REMARKS

Reconsideration and allowance of the subject patent application are respectfully requested.

The specification has been amended to add headings and to correct a minor informality.

Claims 1 and 7 have been amended along the lines kindly suggested by the Examiner in order to address the objections noted on page 2 of the office action.

Claims 1-6 were rejected under 35 U.S.C. Section 101 as allegedly being directed to nonstatutory subject matter. Applicant traverses this rejection. Applicant respectfully submits that claims 1-6 fall within the statutory category of patentable subject matter under 35 U.S.C. Section 101 of a machine. Specifically, these claims define a "system" comprising various means for performing particular functions, and so are clearly directed to a system such as a computer system. To further emphasize this, claim 1 has been amended to refer to a "computer" system.

Withdrawal of the Section 101 rejection of claims 1-6 is respectfully requested.

Claims 1, 3, 7, 9 and 10 were rejected under 35 U.S.C. Section 103(a) as allegedly being made "obvious" by Gryaznov et al. (U.S. Patent No. 7,210,041) in view of Beetz (WO 02/10888).

While not acquiescing in the rejection or in the characterizations of the applied documents in the office action, independent claims 1 and 7 have been amended to clarify the nature of the groups of constituent parts. In particular the groups are defined as each

containing parts of a different type of structural part of the program. This amendment has basis in the disclosure which makes it clear that the groups contain different types of structural parts, for example, comments, variable names, subroutine names, or strings. See, e.g., page 4, lines 4 to 7. New claims 12 and 14 specify examples of types of structural parts.

Conforming amendments have been made to the dependent claims as appropriate.

Finally, new claim 12 has been added claiming the same feature as claim 6 but dependent on method claim 7.

The method of claim 7 relates to scanning files to detect malware. It is particularly concerned with analyzing files containing source code in a given computer language. The method involves obtaining a frequency distribution of characters in groups of constituent parts, which groups comprise parts of a different type of structural part of the program. The frequency distributions are compared with expected ranges as the basis for flagging the file as suspect or not.

Therefore, the method is based on the principle that, in the case of source code, the frequency distribution of characters in different types of structural parts of the program are indicative of whether or not the source code is malware. In general terms, this method is not obvious because no prior art document discloses or suggests that the frequency distribution of characters in respective groups each containing a different type of structural part of source code is indicative of whether or not the source code is malware.

Turning now to the documents relied on in the office action, Gryaznov et al. relates to a system for scanning files (and also data strings) for macro viruses. This is performed using a database of macro virus definitions in data files. Thus Gryaznov et al. relates to a method of scanning a computer file containing source code. In particular, Gryaznov et al. relates to the case of source code which is in a macro language (which is relevant to why the pending claims are not obvious as discussed below).

According to Gryaznov et al., during analysis, a suspect file 27 is parsed into individual tokens. As shown in Fig. 7, the tokens are stored in a parse tree which separates string constants and source code text. In Gryaznov et al., the string constants and source code text are processed to detect a virus. As stated at col. 6, lines 38 to 42, the relevant processing to detect whether a subject file 27 contains a virus is step 87 of Fig. 8, described in more detail in Fig. 9 and at col. 6, line 54 to col. 7, line 33. This processing involves iteratively processing each entry in the parse tree (col. 6, line 67 to col. 7, line 1), i.e., each string and each item of source code text. The processing compares each entry with a virus definition to detect matches with strings and nodes of source code text in the virus definitions (col. 7, lines 5 to 9 and lines 19 to 22). The total number of matches is counted (col. 7, lines 9 to 10 and lines 23 to 24).

Even assuming Gryaznov et al. is viewed as disclosing processing parts and flagging the file as suspect or not on the basis of a comparison, the following claim 7 features are not shown in Gryaznov et al.:

processing each group to count the number of occurrences in that group of characters of a character set to obtain a frequency distribution of characters in that group;

comparing the character frequency distribution of each group with an expected range of frequency distribution; and

flagging the file as suspect or not depending on the result of one or more comparisons by the comparing means.

The office action contends that the features not shown or suggested by Gryaznov et al. are made obvious by Beetz. Applicant traverses this contention, at least in view of the amendments now made.

Beetz relates to a method of scanning files to detect packed executables, i.e., executables in which the viral contents have been subject to compression and need to be expanded before execution. It operates by deriving the frequency distribution of the byte values in the file without unpacking the file and feeding this to a neural network which produces an output indicating the likelihood that the file contains a packed executable. A file detected as being a packed executable is treated as suspicious.

Accordingly, Beetz discloses a specific case of using frequency distributions to detect malware. However, there are significant differences from the claim 7 method. Beetz is concerned with detecting packed executables and derives frequency distributions of byte values in a file. Therefore, Beetz derives a frequency distribution across the entire file and provides the skilled reader with the specific teaching that frequency distributions of byte values in a file are indicative of packed executables which are themselves suspect. The skilled reader is not taught the principle of the claim 7 method

that the frequency distribution of characters in a respective groups each containing a different type of structural part of source code is indicative of whether or not the source code is malware.

Therefore Beetz does disclose or suggest any of the features set out above as lacking in Gryaznov et al. Therefore combining the test applied in Beetz (considering the frequency distribution of byte values across the file) with the test applied in Gryzanov et al. (matching individual strings and items of source code text with strings and items of source code text in virus definitions) does not result in a method within the scope of claim 7.

Moreover, Beetz would not have made it obvious to modify the virus detection technique of Gryaznov et al. to arrive at the claim 7 method. Gryaznov et al. is based on matching of individual strings and items of source code text. Beetz is concerned with derivation and use of a frequency distribution across the file because the frequency distribution of the file can be indicative of a packed file which it itself suspicious. In contrast, and noting in particular the amendments now made, the claim 7 method obtains and uses a frequency distribution of characters in respective groups each containing a different type of structural part of source code.

First, there is no reason why the skilled person reading Beetz would consider deriving a frequency distribution for anything other than the whole file. Beetz is directed towards detection of packed files and it goes against the teachings of Beetz to derive a

frequency distribution for a specific element of the file, such as the "groups" of "parts" as defined in the claims, or indeed any other element of the file.

Second, as Beetz is concerned with detection of packed files, the data of such files is compressed such that the source and "different types of structural parts of the program" are not recognizable, unless and until the file is unpacked (which cannot occur in Beetz because there is *a priori* no knowledge that the file is packed until after the detection). Thus, it would not have been obvious to apply the techniques of Beetz to the "groups" of "parts" as defined in the claims as being structural parts of source code. Conversely, in the context of Beetz, if the file was known to contain source code it would be known not to be packed and hence the technique of using frequency distributions to detect a packed file would be needless.

Third, Gryaznov et al. is concerned with matching of the individual "parts" and not the groups of parts as a whole, so there is no reason why the skilled person reading Beetz would consider deriving a frequency distribution for the specific case of the "groups" of "parts" as defined in the claims.

Claim 1 contains features similar to those of claims 7 and similar arguments are applicable.

Claim 3 depends from claim 1 and claims 9 and 10 each depends from claim 7. These dependent claims patentably distinguish over the applied documents because of their respective dependencies and because of the additional patentable features recited therein.

The Weber et al., Crosbie et al. and Radatti documents are applied in connection with certain dependent claims. These documents do not remedy the deficiencies of the proposed Gryaznov et al.-Beetz combination with respect to the independent claims and, for at least this reason, these dependent claims patentably distinguish over the various proposed combinations of documents.

New claims 12-16 have been added. Applicant respectfully submits that these claims find support in the original disclosure and the Examiner is invited to independently confirm that this is the case.

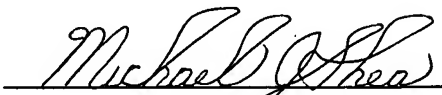
New claims 12-14 are discussed above.

New claim 15 is for a computer-readable medium having stored thereon instructions for causing a computer to carry out a method like that of claim 7. Consequently, this claim patentably distinguishes over the applied documents for the reasons discussed above with respect to claim 7. Claim 16 refers to claim 15.

The pending claims are believed to patentably distinguish over the applied documents and favorable office action is respectfully requested.

Respectfully submitted,

NIXON & VANDERHYE P.C.

By: 

Michael J. Shea
Reg. No. 34,725

MJS:mjs
901 North Glebe Road, 11th Floor
Arlington, VA 22203-1808
Telephone: (703) 816-4000
Facsimile: (703) 816-4100